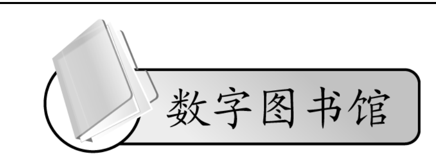


我国教育科研机构知识库建设现状调查与分析

姚晓霞 [1,2]　聂　华 [1]　顾立平 [3]　张冬荣 [3]　吴　越 [1]　韦成府 [1]
[1](北京大学图书馆　北京 100871)
[2](中国高等教育文献保障系统(CALIS)管理中心　北京 100871)
[3](中国科学院文献情报中心　北京 100190)

摘要:【**目的**】为求更好地了解中国机构知识库的发展现况, 以及提供未来发展战略的建议。【**方法**】针对 130 个样本, 进行线上问卷调查。【**结果**】我国机构知识库在内容存储、系统平台、管理政策等三方面各有不同程度的发展。【**局限**】总体调查规模有限, 调查结果的分析聚焦在已经建设机构知识库的单位。【**结论**】调查报告全面总结和分析我国教育科研机构知识库的建设现状, 建议未来进行机构知识库最佳案例的汇编。

关键词: 机构知识库　中国教育科研机构　开放获取
分类号: G250

1　研究背景与调查需求

开放获取学术资源日益增长, 成为学术信息资源管理的新兴课题之一[1], 对图书馆和信息中心而言, 面对开放学术资源环境, 需要重新审视和认识到“什么是图书馆馆藏、如何建设图书馆馆藏、如何组织开放资源、如何支持开放出版、如何对开放资源进行长期保存”等一系列问题[2]。其中, 机构知识库(Institutional Repository, IR)作为收集、存储与传播某个或多个学术机构(例如大学、研究所、图书馆、博物馆等)所创造的、可供机构内外用户共享的学术文献数据库, 具有协助教育科研机构知识资产保存和知识管理的重要作用。

为此, 国内已有许多实践工作, 例如大陆高校推动机构知识库的实践[3]、台湾高校发展机构知识库的介绍[4]、香港高校发展机构知识库的个案[5]、科研院所的机构知识库建设[6], 以及研究型图书馆推进开放获取的实践总结[7]等。

为了促进我国机构知识库的广泛建设与健全发展, 中国机构知识库推进工作组于 2012 年 10 月成立[8]; 在其 2013 年度工作计划中, 包括“中国机构知识库学术研讨会”[9]等推广活动和支持其他单位的工作任务。其中一项主要任务是开展针对我国机构知识库建设现况的调查, 以求完整地掌握目前发展现况, 以便从事长远的机构知识库发展实施战略以及执行方案。

2　研究目标与调查问题

IR 承担着保存机构知识资产和支持知识开放共享的重要职责, 是新型学术信息交流体系和教育科研知识基础设施的有机组成部分。当前, 已有不少国内研究利用 Open DOAR(The Directory of Open Access Repositories)、ROAR(Registry of Open Access Reposi-

收稿日期: 2014-03-24
收修改稿日期: 2014-04-04
*本文系“2013 中国机构知识库”学术研讨会会议论文。

tories) 和 Ranking Web of Repositories(The Spanish National Research Council)等平台中的各项数据对中国大陆，港台地区甚至全球的机构知识库建设现状进行调研和分析。例如，赵莉娜等[10]通过对 ROAR 网站中统计数据信息的调查分析，介绍了国外机构知识库目前的建设情况，重点对我国机构知识库建设情况进行了分析并提出：争取国家政策与项目支持、明确高校图书馆建设机构知识库的引领职责、建立机构知识库建设的动力机制、建立学科馆员协助存储机制、加强技术创新努力开发适合中国用户信息需求的机构知识库系统软件、建立机构知识库联盟等建议。唐义等[11]对港台地区学术机构知识库的建设状况进行整体调查，发现港台地区普遍基于低成本的 DSpace 开发、实行分散建置与集中呈现的发展、教育主管部门提供政策和资金支持。中国高等教育文献保障系统(CALIS)曾在我国高校范围内组织了问卷调查[12]。纪元霞[13]利用 ROAR 数据对国内机构知识库软件平台的采用情况进行了调研。袁顺波等[14]基于 OpenDOAR 数据对全球机构知识库的建设现状进行了调研和分析。王婉[15]对比香港和大陆地区高校机构知识库的建设现状并提出建议。

考虑到我国机构知识库建设尚处于初级阶段，有众多机构库虽然已经投入使用但还拘于包括语言等各种原因尚未在国际注册网站中注册，还有许多机构库正在建设之中；因此，中国机构知识库推进工作组决定组织全国性调查工作，以了解和把握我国机构知识库建设的现状和问题、战略和措施、挑战和困难等，为机构知识库的未来发展提供思考和借鉴。

因此，本次调查针对存缴什么内容、按照什么方式存缴、存缴内容的类型、如何进行存储、如何针对存缴的内容开展服务、未来如何发展等方面重点展开调查。内容包括：内容存缴方式、数据存缴的操作方式、存缴内容类型与载体格式、存缴激励措施、政策内容、开放获取时滞期、IR 的存储量与用户量、软件平台、服务功能、宣传推广、构建目的及当前的问题等。

3 调查设计与实施过程

3.1 问卷设计过程

根据《中国机构知识库 2013 年度工作任务计划》设计问卷初稿，在 CALIS 三期机构知识库建设与推广子项目的调查基础之上，进行全面修改、补充扩展、删减；征求工作组专家意见后，选取用户预填写调查问卷，反复修订问卷 5 次后确定问卷终稿。

3.2 问卷调查规划

调查样本是具有代表已建或者待建设机构知识库的教育科研机构中的图书馆员，即一个机构一个样本。调查对象分为两大系统：科研系统和高校系统图书馆。科研系统图书馆的调查对象由中国科学院文献情报中心牵头组织；高校系统图书馆的调查对象由北京大学牵头组织。

3.3 问卷发放回收的实施情况

调查采用在线问卷系统，分为 A 卷和 B 卷：A 卷针对已经建设 IR 的教育科研机构，B 卷针对拟建设 IR 的教育科研机构。两份问卷的题目设置完全一致，但 A 卷要求所有题目为必答，B 卷中除基本信息题为必答之外，其他题目均为选答。调查于 2013 年 5 月 15 日至 2013 年 7 月 10 日进行，共收回问卷 130 份，其中有效问卷 101 份，无效问卷 29 份。在 101 家教育科研机构中，有 26 所为中国科学院科研机构，75 所为国内高等院校。

3.4 问卷反馈情况

根据调查结果，有 40 家教育科研机构反馈已经建设机构知识库，61 家未建设或待建设。

在 40 家已建设 IR 的教育科研机构中，有 21 所来自中国科学院科研机构，19 所来自国内高等院校。61 家未建设 IR 或待建设的教育科研机构，有 5 所来自中国科学院科研机构，56 所来自国内高等院校。

4 调查结果

4.1 IR 内容的存缴以数据迁移复制的比例较高

在存缴方式的 5 个选项(A.个人自行存缴；B.个人提交管理部门，管理部门人员进行存缴；C.个人提交管理部门，管理部门提交 IR 管理者；D.个人提交 IR 管理者；E.从其他平台进行数据迁移或者数据复制)中，从其他平台进行“数据迁移或复制”的比例最高，如图 1 所示。

4.2 IR 内容存缴的操作方式以个人上网提交和图书馆批量导入的情况居多

调查结果显示，多种操作途径并进，对 A 卷的进一步分析看到“在 IR 网页上提交数据信息”(主要是个人提交)和“在 IR 后台，以程序命令、批量导入数据”(即由图书馆等中介提交内容)这两种方式被采用的比例最高，各有将近 80%的机构采用，如图 2 所示。

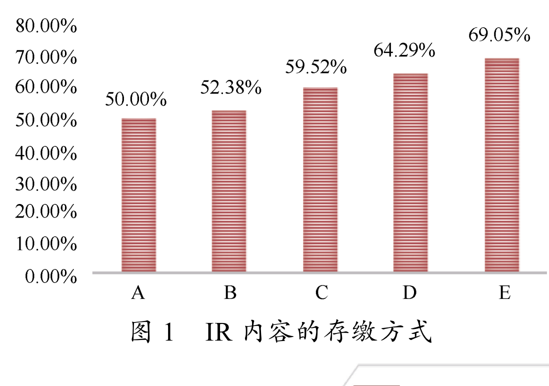


图 1　IR 内容的存缴方式

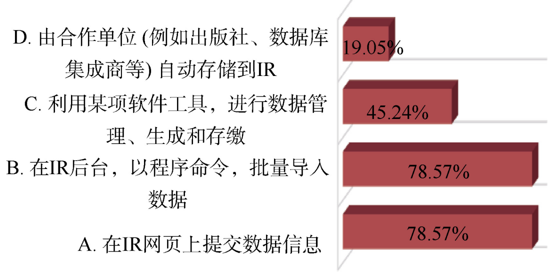


图 2　IR 内容存缴的操作方式

4.3　IR 存缴内容的类型以“期刊论文、会议论文、学位论文”最多

调查显示各个机构知识库均认可存缴内容的类型多样化，其中对“期刊论文、会议论文、学位论文”最为重视，达到被调查机构的 95%以上，如图 3 所示:

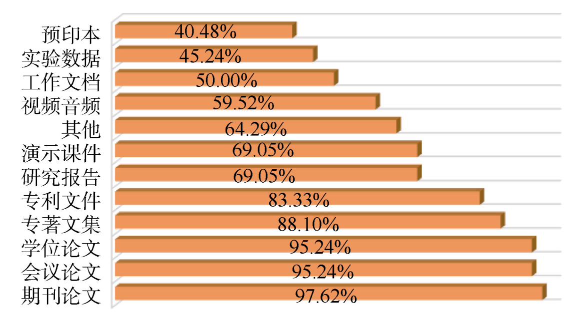


图 3　IR 存缴内容类型

4.4　IR 存缴的载体格式主要支持“文本、表格、图片和课件”

存缴内容的载体格式也是多样的，对当前主流的 4 种载体格式(文本、表格、图片和课件)支持达到 80%以上，如图 4 所示。

4.5　具有 IR 存缴激励措施的机构不多，少数采取“强制要求”和“考核评估”等方式

调查显示，6 种激励措施采用比例都不高。然而，有 35%以上的被调查单位具有“强制要求”和/或“考核评估”的措施，如图 5 所示。

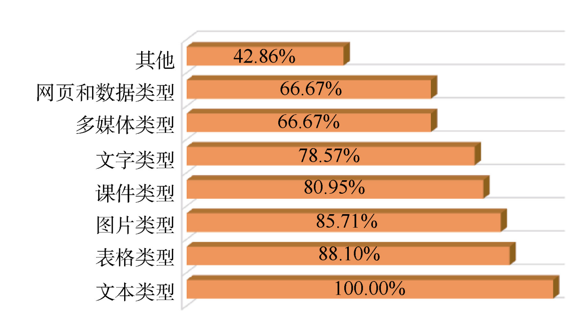


图 4　IR 载体格式

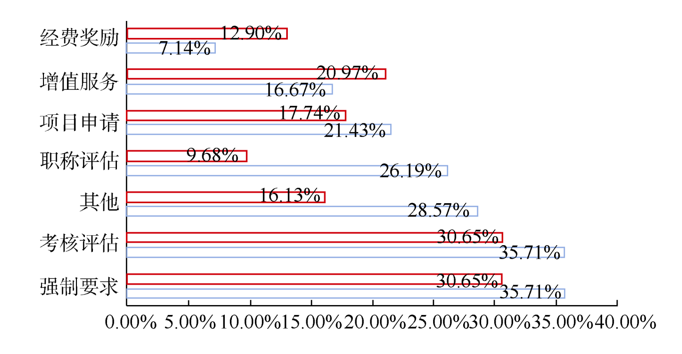


图 5　IR 存缴激励措施

4.6　将近一半的 IR 具有明确的存储与传播政策

调查显示已建 IR 的知识库有 52%具备严格统一的存缴政策，有 36%的机构并没有明确的 IR 存缴政策；对拟建设 IR 的机构调研显示，有 52%目前没有考虑制定明确的存缴政策，如图 6 所示:

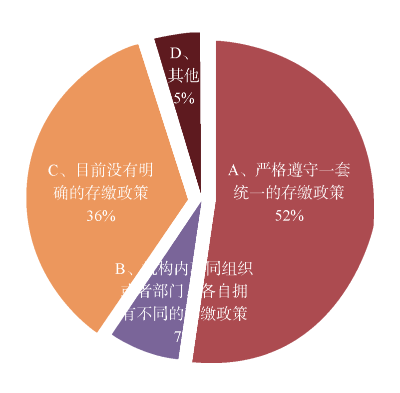


图 6　IR 存缴政策

在 IR 开放服务政策上，已建知识库的回复中，有 59%的 IR 严格遵守一套统一的开放服务政策，而有 29%没有明确的开放服务政策；正在建设 IR 的调查回复中，表达了对统一的开放服务政策的关注，但在开放服务政策的选择方面尚不明确，如图 7 所示。

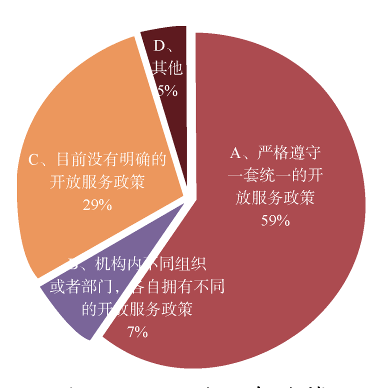


图 7　IR 开放服务政策

4.7　多数已建 IR 采取机构统一规定存缴内容开放程度的管理方式

调查显示，已建 IR 的机构有 76%采取“机构统一规定存缴内容的开放程度”的方式；拟建 IR 的机构仅有不到一半的机构对于存缴内容的开放程度有统一的规定，如图 8 所示：

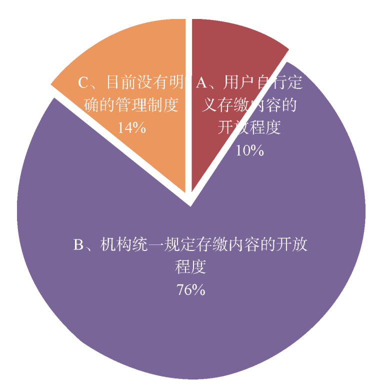


图 8　IR 存储内容实施开放共享的管理方式

在开放获取时滞期方面，选择“存缴后立即实施开放共享”的机构超过一半，如图 9 所示：

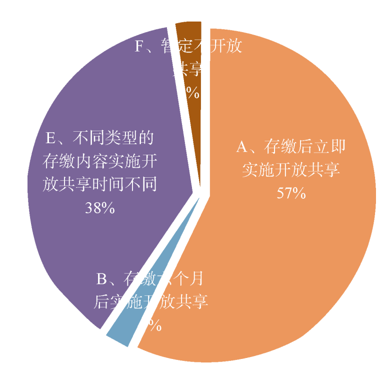


图 9　IR 实施开放共享的时间

在 IR 浏览、检索、下载的权限控制上，根据调查结果，被调查的 IR 中仅有极少数的机构在开放共享方面采取不开放的措施，如图 10 所示：

图 10　IR 浏览、检索和下载的权限控制

4.8　已建 IR 的存储量已有较大规模

根据调查结果，已建 IR 的机构中，元数据条目达 10 万条以上的机构有 14%；另有 38%的 IR 元数据量也已具有一定的规模，在 1 万条-10 万条之间，如图 11 所示：

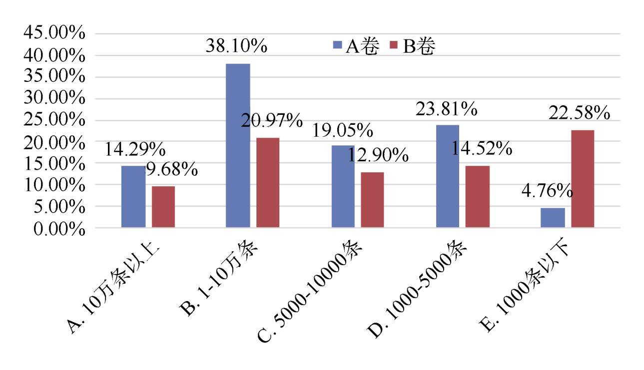


图 11　IR 数据量统计

在全文总量方面，已建 IR 的机构中，有 7%的 IR 全文总量达到 10 万条以上，而有 43%的全文总量在 1 万条以上，如图 12 所示：

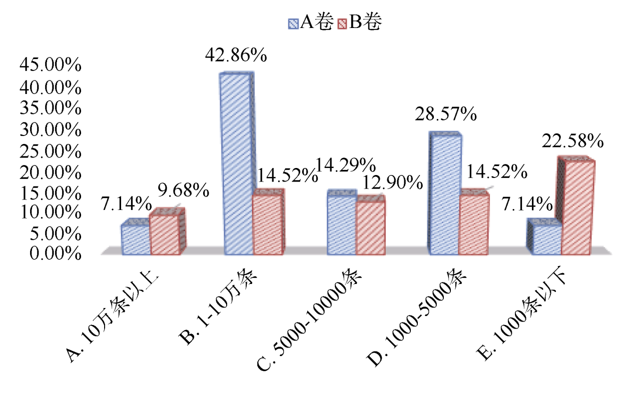


图 12　IR 全文量统计

在 IR 当前用户注册数量方面，在已经建设的 IR 方面，有 7%具有 1 万以上的注册用户；有 10%的 IR 具有 1 000-10 000 注册用户的规模，有 83%的 IR 注册用户在 1 000 人以下，如图 13 所示。

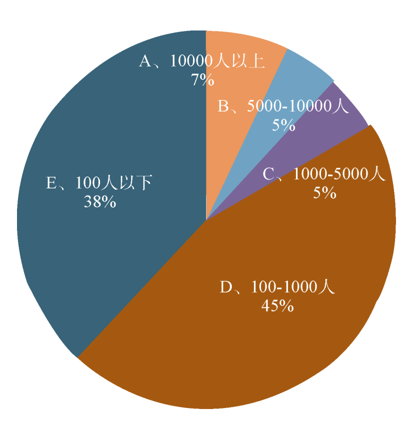


图 13　IR 用户注册数量统计

4.9　建设 IR 平台的方式多样，以其他机构协助建设的情况稍多

在被调查机构当中，建设 IR 平台的方式有多种；其中，以协助建设方式最为突出，开源软件的应用成为主流，如图 14 所示：

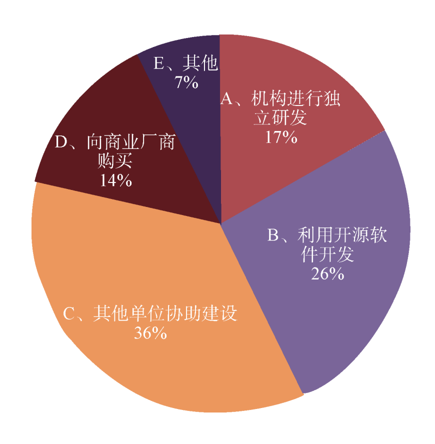


图 14　IR 平台的建设方式

在开源工具的利用上，已建 IR 的回复中，有 62% 的 IR 利用开源软件 DSpace 进行 IR 系统的开发，也有 12%的被调查机构选择利用其他开源工具，另有 26% 的 IR 没有利用开源软件，如图 15 所示。

4.10　多数 IR 选择内容集中管理

在内容组织管理方式上，无论 A 卷还是 B 卷，元数据和全文集中管理是被采用的最多的内容组织管理方式，比例分别达 81%和 66%等；另有少数 IR 选择采用“元数据集中管理，全文由机构内不同部门和组织各自管理”或“IR 元数据和全文由机构内不同部门和组织各自管理”的方式，如图 16 所示。

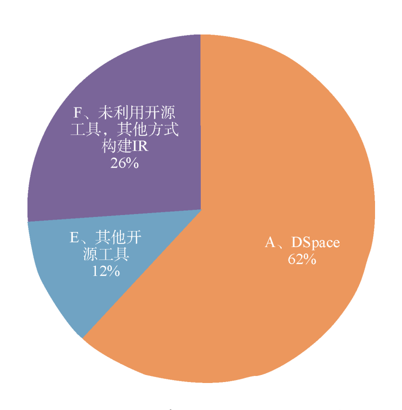


图 15　IR 对开源工具的利用

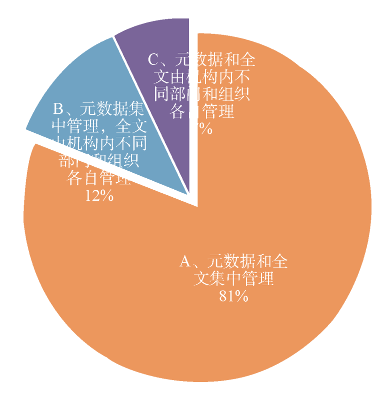


图 16　IR 内容组织管理方式

4.11　新型服务功能视需要而建

调查显示，已建 IR 的知识库多数具有“主页检索功能、高级检索功能和显示信息资源类型”等服务功能，比例在 90%以上，如图 17 所示。但是，一些新型功能尚待广泛应用。

4.12　学科馆员走访作为主要的 IR 宣传推广方式

如图 18 所示，在宣传推广范围方面，A 卷中的被调查机构最常使用到的宣传推广方式是“在机构单位范围内推广”，被采用的比例达 83.33%。另外采用“全国范围内推广(参加国内会议和 IR 联盟)”的机构也达到被调查 IR 的近 60%。国际范围内推广的采用比例最低，但也是提高 IR 影响力，逐步和国际接轨所应当重视的一种宣传推广方式。

在宣传推广方式上，“学科馆员走访”是被采用的最多的 IR 宣传推广方式；而“开展讲座培训”和“发布新闻”也是被普遍采用的方式，如图 19 所示。

4.13　建设 IR 的主要目的是保存知识资产和扩大学术传播影响力

根据调查结果，A 卷和 B 卷在建设 IR 的目的上具有高度的一致性，对 IR 建设目的中“完善保存机构

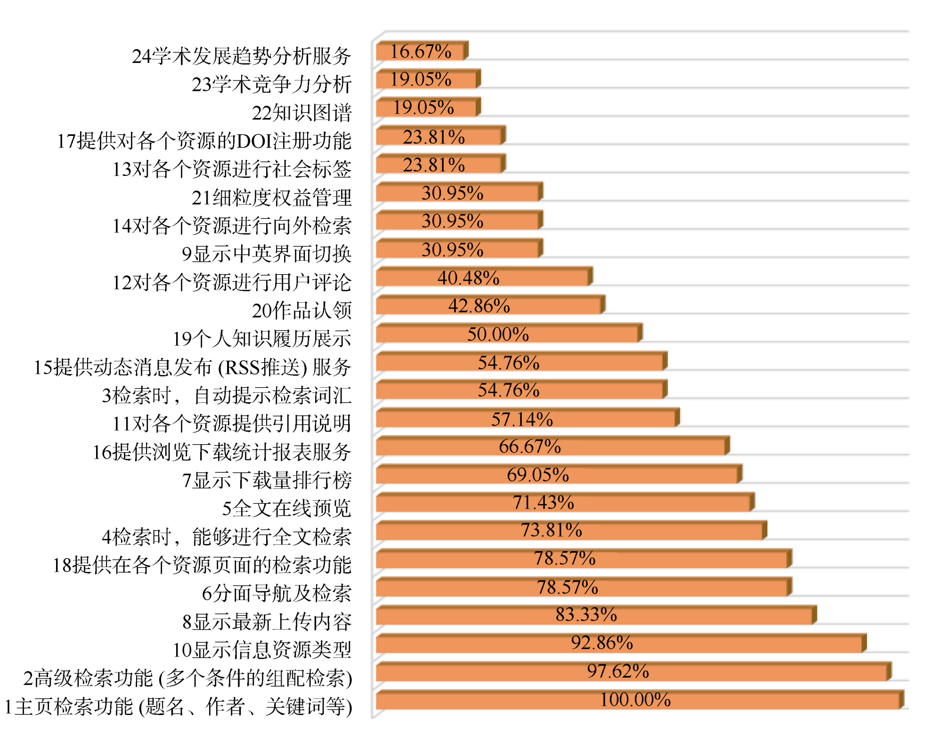


图 17　IR 服务功能

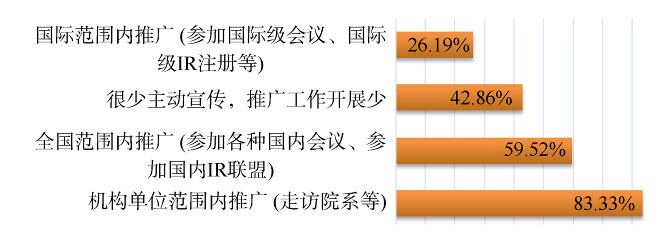


图 18　IR 宣传推广范围

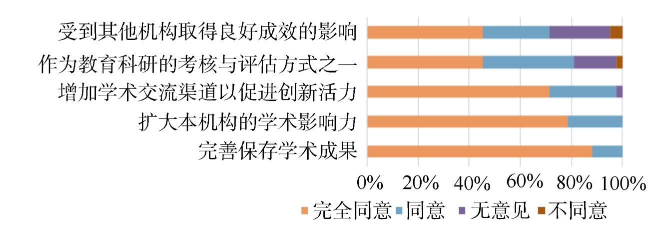


图 20　IR 建设目的

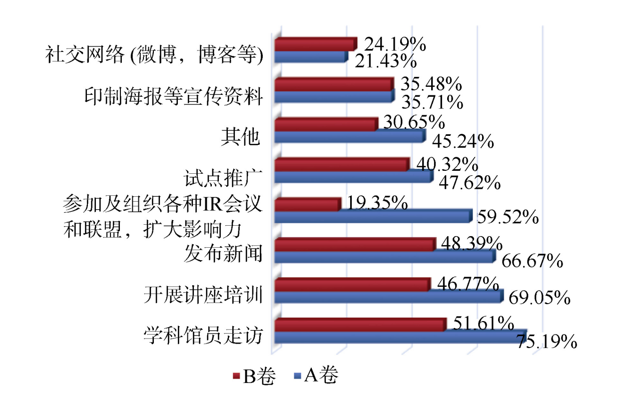


图 19　IR 宣传推广方式

的学术成果”和“扩大本机构的学术影响力”的支持度最高，也认同 IR“增加学术交流渠道以促进创新活力”和“作为教育科研的考核和评估方式”的作用，如图 20 所示。

4.14　经费资助和政策支持是下一步需要解决的问题

对于调查问卷中预设的构建 IR 下一步需要解决的问题，包括“经费、政策(版权开放和保护)支持、人力、内容、宣传、参与积极性”等，90%左右的调查机构都认为调查问卷中列出的问题是本机构需要解决的问题，60%的机构认为这些问题迫切需要解决，如图 21 所示：

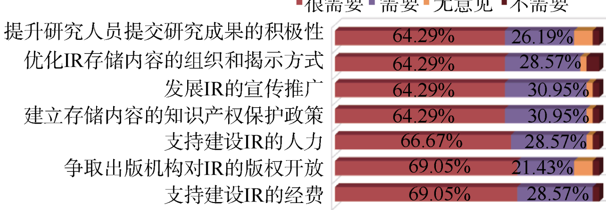


图 21　IR 下一步需要解决的问题

5　结论与讨论

(1) 我国机构知识库存储进入快速增长期

从对已建的数据量、用户量和访问量等方面的调查结果来看，当前已有一定机构知识库的建设初具规模，IR 元数据量达到 10 万条以上，全文总量达到 10 万条以上，有 1 万人以上的注册用户，IR 月均访问量在 1 万以上。这说明，我国机构知识库已经开始彰显实际的服务效益。

(2) 图书馆应当争取合作存储的方式

在 IR 内容和元数据存缴的操作方式上，网页提交(个人)和后台批量导入数据(中介)采用比例最高。而“合作单位自动存储”采用比例最少，可能的原因是合作单位(如出版社、数据库集成商等)目前对 IR 的认识和积极性仍然不够，IR 建设过程中图书馆需要加强与出版社、内容商等的合作，探索合作和参与的双赢的、适用的方式和途径。

(3) 图书馆应积极争取教育科研机构层面的激励政策

在调查中，有关存缴激励措施的 7 种方式(包括其他)采用比例偏低，都在 40%以下。图书馆和信息中心等，应该相互借鉴经验，争取在教育科研机构中实施激励 IR 存缴与传播的措施。一方面完善 IR 的存缴和传播服务政策，另一方面明确管理和考核等激励措施，其中包括灵活地为存储者提供“不同类型的存缴内容实施开放共享的时间不同”的选择等，最后，根据机构政策制定或完善 IR 管理政策。

(4) 拟建 IR 的机构可以考虑借鉴已有 IR 机构的经验

在平台构建方面，国内大部分机构对基于 DSpace 开发建设 IR 十分认可，并已经积累了一定的经验，拟建 IR 的机构可以充分地借鉴已有的经验，采用开源软件进行 IR 建设。这些实践经验可以在未来通过案例汇编的工作，予以分享。

另一方面，当前各个机构库提供的服务功能十分丰富，彼此可以根据需要而相互借鉴，例如个人知识履历展示、作品认领、细粒度权益管理、知识图谱等。

(5) 建议未来汇编《机构知识库建设实践案例》等 IR 建设指南

我国教育科研机构对于机构知识库的建设目标有较高的共识，大部分机构建设 IR 的目的是为了完善保存机构的学术成果并扩大机构的学术影响力。但是当前在 IR 建设中也存在一定的问题，包括：研究人员提交成果的积极性、内容的组织方式、宣传推广方式、知识产权政策、人力和经费的评估、出版商对版权控制等。这些问题不应成为阻挠 IR 建设的障碍；相反，未来应当汇编我国的《机构知识库建设实践案例》，以简单易懂的方式，传播 IR 建设中使用的各种系统技术、政策法规、宣传推广方式等操作指南，为健全 IR 发展而努力。

本次调查相对完整地呈现了我国高校和科研院所机构知识库的建设情况，通过对调查结果的分析，提出了上述 5 项建设性策略，提供业界参考。

(致谢：中国学术期刊对机构知识库存缴政策的调查是在“中国机构知识库推进工作组”(http://chinair.csdl.ac.cn/node/)的规划、组织和指导之下实施。共有 101 家科研教育机构积极参与了调查，在此一并感谢。)

参考文献：

[1] 张晓林，李麟，刘细文，等. 开放获取学术信息资源：逼近“主流化”转折点[J]. 图书情报工作，2012，56(9)：42-47. (Zhang Xiaolin, Li Lin, Liu Xiwen, et al. Open Access Scholarly Resources: Getting Close to Being the Main-Stream[J]. Library and Information Service, 2012, 56(9): 42-47.)

[2] 张晓林，曾燕，李麟，等. 开放学术信息资源环境的挑战及其应对策略[J]. 图书情报工作，2012，56(19)：5-12，17. (Zhang Xiaolin, Zeng Yan, Li Lin, et al. Analysis of Challenges and Responses to the Imminent Open Access World[J]. Library and Information Service, 2012, 56(19): 5-12, 17.)

[3] 聂华，韦成府，崔海媛. CALIS 机构知识库：建设与推广、反思与展望[J]. 中国图书馆学报，2013，39(2)：46-52.(Nie Hua, Wei Chengfu, Cui Haiyuan. CALIS Institutional Repository: Construction and Promotion, Reflection and Prospects[J]. Journal of Library Science in China, 2013, 39(2): 46-52.)

[4] 包心萍，刘兹恒. 台湾高校图书馆机构知识库发展研究[J]. 大学图书馆学报，2012，30(4)：61-65. (Bao Xinping, Liu Ziheng. Study on the Development of Institutional Repositories of the University Libraries in Taiwan of China[J]. Journal of Academic Libraries, 2012, 30(4):61-65.)

[5] 顾立平. 提升机构知识库作用的实践——多元学科教育资

源的集成与再利用[J]. 图书情报工作网刊, 2011(2): 55-82. (Ku Liping.Practice on Enhance the Role of Institutional Repositories——Multidisciplinary Education Resource Integration and Re-Cycle Use[J]. Library and Information Service Online, 2011(2): 55-82.)

[6] 张冬荣, 祝忠明, 李麟, 等. 中国科学院机构知识库建设推广与服务[J]. 图书情报工作, 2013, 57(1): 20-25. (Zhang Dongrong, Zhu Zhongming, Li Lin, et al. Construction, Promotion and Service of CAS IRs[J]. Library and Information Service, 2013, 57(1): 20-25.)

[7] 张晓林, 刘细文, 李麟, 等. 研究图书馆推进开放获取的战略与实践——以国家科学图书馆为例[J]. 图书情报工作, 2013, 57(1): 15-19, 48. (Zhang Xiaolin, Liu Xiwen, Li Lin, et al. The Strategies and Practices of Research Library to Support Open Access ——Taking National Science Library as an Example[J]. Library and Information Service, 2013, 57(1): 15-19, 48.)

[8] 中国科学院国家科学图书馆.中国机构知识库推进工作组成立[OL]. [2012-11-07]. http://www.las.cas.cn/xwzx/zhxw/201211/t20121107_3678334.html.(National Science Library, Chinese Academy of Sciences. China Institutional Repository Implementation Group (CIRIG) has Established [OL]. [2012-11-07].http://www.las.cas.cn/xwzx/zhxw/201211/t20121107_3678334.html.)

[9] 中国机构知识库推进工作组.国内首次机构知识库会议成功举办, 推动科技信息开放获取[OL]. [2013-09-30]. http://2013chinair.csp.escience.cn/dct/page/1. (China Institutional Repository Implementation Group. Success of the First Institutional Repository Conference in China, Puch on Open Access of S&T Information [OL]. [2013-09-30]. http://2013chinair.csp.escience.cn/dct/page/1.)

[10] 赵莉娜, 徐春艳. 国内外机构知识库建设情况调查研究[J]. 图书馆学研究, 2011(6): 38-43. (Zhao Li'na, Xu Chunyan. Research on the Institutional Repositories[J]. Library Studies, 2011(6): 38-43.)

[11] 唐义, 肖希明. 港台学术机构知识库的调查与分析[J]. 图书馆论坛, 2011, 31(4): 38-41, 9. (Tang Yi, Xiao Ximing. Survey and Analysis to the Institutional Repository in Hong Kong and Taiwan[J]. Library Forum, 2011, 31(4): 38-41, 9.)

[12] Zhao Y C, Yao X X, Wei C F. Academic Institutional Repositories in China: A Survey of CALIS Member Libraries[J]. Chinese Journal of Library and Information Science, 2012, 5(2): 18-32.

[13] 纪元霞. 国内机构库软件平台调研[J]. 图书情报工作网刊, 2012(1): 31-39. (Ji Yuanxia. Survey of Repositories Software in China Mainland[J]. Library and Information Service Online, 2012(1): 31-39.)

[14] 袁顺波, 董文鸳, 刘艳华. 机构库实践建设现状研究——基于 OpenDOAR 的调研与分析[J]. 图书馆, 2012(5): 67-70. (Yuan Shunbo, Dong Wenyuan, Liu Yanhua. Research on the Repositories Building Practice——A Study Based on the OpenDOAR[J]. Library, 2012(5): 67-70.)

[15] 王婉. 香港地区大学机构知识库的调查与分析[J]. 农业图书情报学刊, 2012, 24(7): 101-104. (Wang Wan. Survey and Analysis of the Institutional Repository of Universities in Hong Kong [J]. Journal of Library and Information Sciences in Agriculture, 2012, 24(7): 101-104.)

作者贡献声明:

姚晓霞, 聂华, 顾立平: 提出调查思路, 组织设计调查方案;
聂华, 顾立平, 姚晓霞, 张冬荣, 吴越, 韦成府: 调查问卷的设计与修改;
聂华, 姚晓霞, 张冬荣, 顾立平: 组织实施调查;
韦成府: 在线调查平台设计和调查数据收集;
吴越, 姚晓霞, 韦成府: 问卷数据分析与论文初稿;
顾立平, 聂华: 论文最终版本的修订。

(通讯作者: 聂华 E-mail: hnie@lib.pku.edu.cn)

Survey and Analysis of the Status on Institutional Repository in Chinese Academic and Research Libraries

Yao Xiaoxia[1,2]　Nie Hua[1]　Ku Liping[3]　Zhang Dongrong[3]　Wu Yue[1]　Wei Chengfu[1]
[1](Peking University Library, Beijing 100871, China)
[2](CALIS Administrative Center, Beijing 100871, China)
[3](National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: [**Objective**] This survey is for better understanding of the situation of Institutional Repositories in China mainland and to make the recommendation of the future development strategy. [**Methods**] 130 samples are selected to do the online questionnaire survey. [**Results**] There are many repositories in China that developed their content deposit, system platform and management policy on different levels. [**Limitations**] The survey scale is limited, and the analysis focuses on the institutions which have developed repositories. [**Conclusions**] The investigation report comprehensively summarizes and analyzes the IR current situation on Institutional Repository in Chinese Academic and Research Libraries. The recommendation is to edit the best practice of Institutional Repositories in future.

Keywords: Institutional Repository　Chinese education research institution　Open access